# FREE PROBABILITY THEORY

ROLAND SPEICHER

## Lecture 3
## Freeness and Random Matrices

Now we come to one of the most important and inspiring realizations of freeness. Up to now we have realized free random variables by generators of the free group factors or by creation and annihilation operators on full Fock spaces. This was not very surprising because our definition of freeness was modelled according to such situations. But there are objects from a quite different mathematical universe which are also free (at least asymptotically) - namely special *random matrices*. Random matrices are matrices whose entries are classical random variables, and the most important class of random matrices are the so-called Gaussian random matrices where the entries are classical Gaussian random variables. So, before we talk about random matrices, we should recall the basic properties of Gaussian random variables.

### 3.1. **Moments of Gaussian random variables.**

**Definition 1.** 1) A family of classical real-valued random variables $x_1, \ldots, x_n$ is a *(centered) Gaussian family* if its joint density is of a Gaussian form, i.e.,

$$(2\pi)^{-n/2}(\det C)^{-1/2}\exp\{-\frac{1}{2}\langle x, C^{-1}x\rangle\},$$

where $x = (x_1, \ldots, x_n)$ and $C$ is a non-singular $n \times n$-matrix (which determines the covariance of the variables). This is equivalent to saying that the characteristic function of $x$ is of the form

$$E[e^{i\langle t,x\rangle}] = \exp\{-\frac{1}{2}\langle t, Ct\rangle\}.$$

This latter form also makes sense if $C$ is singular.
2) A family of classical complex-valued random variables $z_1, \ldots, z_n$ is

Research supported by a Discovery Grant and a Leadership Support Inititative Award from the Natural Sciences and Engineering Research Council of Canada and by a Premier's Research Excellence Award from the Province of Ontario.

Lectures at the "Summer School in Operator Algebras", Ottawa, June 2005.

a *Gaussian family* if the collection of their real and imaginary parts $\Re z_1, \Im z_1, \ldots, \Re z_n, \Im z_n$ is a Gaussian family.

The prescription of the density allows to calculate all moments in Gaussian variables by calculating integrals with respect to the above density. Clearly, all mixed moments are determined by the covariance matrix $C$, which is just the information about all second moments of our variables.

A much better way of describing such a Gaussian family (at least for our purposes) is, however, by saying very concretely how higher moments are determined in terms of second moments. This is a very nice combinatorial description, which usually goes under the name of *Wick formula*: We have for all $k \in \mathbb{N}$ and all $1 \leq i(1), \ldots, i(k) \leq n$ that

$$E[x_{i(1)} \cdots x_{i(k)}] = \sum_{\pi \in \mathcal{P}_2(k)} \prod_{(r,s) \in \pi} E[x_{i(r)} x_{i(s)}].$$

Here, $\mathcal{P}_2$ denotes the set of all pairings of the set $\{1, \ldots, k\}$. Of course, if $k$ is odd, there is no such pairing at all and an odd moment of a Gaussian family has to vanish.

Note that this Wick formula is very similar to the formula for moments of semi-circular elements; in the latter case we just have non-crossing pairings instead of all pairings. This is one justification of the claim that semi-circular elements are the free counterparts of Gaussian random variables.

It's probably the best to illustrate the above Wick formula by a few examples. For $k = 2$, it's just the tautology

$$E[x_{i(1)} x_{i(2)}] = E[x_{i(1)} x_{i(2)}],$$

whereas for $k = 4$ it says

$$\begin{aligned} E[x_{i(1)} x_{i(2)} x_{i(3)} x_{i(4)}] &= E[x_{i(1)} x_{i(2)}] \cdot E[x_{i(3)} x_{i(4)}] \\ &+ E[x_{i(1)} x_{i(3)}] \cdot E[x_{i(2)} x_{i(4)}] \\ &+ E[x_{i(1)} x_{i(4)}] \cdot E[x_{i(2)} x_{i(3)}] \end{aligned}$$

And because it's so nice, here is also the case $k = 6$:

$$\begin{aligned} E[x_{i(1)} x_{i(2)} x_{i(3)} x_{i(4)} x_{i(5)} x_{i(6)}] &= E[x_{i(1)} x_{i(2)}] \cdot E[x_{i(3)} x_{i(4)}] \cdot E[x_{i(5)} x_{i(6)}] \\ &+ E[x_{i(1)} x_{i(2)}] \cdot E[x_{i(3)} x_{i(5)}] \cdot E[x_{i(4)} x_{i(6)}] \\ &+ E[x_{i(1)} x_{i(2)}] \cdot E[x_{i(3)} x_{i(6)}] \cdot E[x_{i(4)} x_{i(5)}] \\ &+ E[x_{i(1)} x_{i(3)}] \cdot E[x_{i(2)} x_{i(4)}] \cdot E[x_{i(5)} x_{i(6)}] \\ &+ E[x_{i(1)} x_{i(3)}] \cdot E[x_{i(2)} x_{i(5)}] \cdot E[x_{i(4)} x_{i(6)}] \\ &+ E[x_{i(1)} x_{i(3)}] \cdot E[x_{i(2)} x_{i(6)}] \cdot E[x_{i(4)} x_{i(5)}] \end{aligned}$$

$$
\begin{aligned}
&+ E[x_{i(1)}x_{i(4)}] \cdot E[x_{i(2)}x_{i(3)}] \cdot E[x_{i(5)}x_{i(6)}] \\
&+ E[x_{i(1)}x_{i(4)}] \cdot E[x_{i(2)}x_{i(5)}] \cdot E[x_{i(3)}x_{i(6)}] \\
&+ E[x_{i(1)}x_{i(4)}] \cdot E[x_{i(2)}x_{i(6)}] \cdot E[x_{i(3)}x_{i(5)}] \\
&+ E[x_{i(1)}x_{i(5)}] \cdot E[x_{i(2)}x_{i(3)}] \cdot E[x_{i(4)}x_{i(6)}] \\
&+ E[x_{i(1)}x_{i(5)}] \cdot E[x_{i(2)}x_{i(4)}] \cdot E[x_{i(3)}x_{i(6)}] \\
&+ E[x_{i(1)}x_{i(5)}] \cdot E[x_{i(2)}x_{i(6)}] \cdot E[x_{i(3)}x_{i(4)}] \\
&+ E[x_{i(1)}x_{i(6)}] \cdot E[x_{i(2)}x_{i(3)}] \cdot E[x_{i(4)}x_{i(5)}] \\
&+ E[x_{i(1)}x_{i(6)}] \cdot E[x_{i(2)}x_{i(4)}] \cdot E[x_{i(3)}x_{i(5)}] \\
&+ E[x_{i(1)}x_{i(6)}] \cdot E[x_{i(2)}x_{i(5)}] \cdot E[x_{i(3)}x_{i(4)}]
\end{aligned}
$$

Note that this Wick formula gives us the even moments of one Gaussian random variable $x$ in the form

$$E[x^{2m}] = \#\mathcal{P}_2(2m) \cdot \sigma^{2m},$$

where $\sigma^2 = E[x^2]$. Clearly, the number of pairings of $2m$ elements is given by $(2m-1) \cdot (2m-3) \cdots 5 \cdot 3 \cdot 1$ (which one usually denotes by $(2m-1)!!$). You might like to check that indeed

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} t^{2m} \exp\{-\frac{t^2}{2\sigma^2}\} dt = (2m-1)!! \cdot \sigma^{2m}.$$

In many cases it is quite useful to have this combinatorial interpretation for the moments of Gaussians.

If the covariance matrix $C$ is diagonal then $E[x_i x_j] = 0$ for $i \neq j$, and this just means that the $x$'s are independent. So in this special case we can describe the Gaussian family as consisting of independent Gaussian random variables. By diagonalizing $C$, one sees that a Gaussian family is given by linear combinations of independent Gaussian random variables.

Strictly speaking, the above Wick formula was for real-valued Gaussian variables $x_1, \ldots, x_n$, however, the nice feature is that it remains the same for complex-valued Gaussian variables $z_1, \ldots, z_n$, not only by replacing the $x_i$ by $\Re z_i$ or $\Im z_i$, but also by replacing them by $z_i$ or $\bar{z}_i$ (of course, this follows from the multilinear structure of the formula) This is important for us because the entries of our Gaussian random matrices will be complex-valued.

3.2. **Random matrices in general.** Let us now talk about random matrices. We start here by being very general in order to fix the language and the frame, but then we will look very quickly on the prominent example of Gaussian random matrices.

An $N \times N$-random matrix is an $N \times N$-matrix whose entries are classical random variables, i.e.,

$$A = (a_{ij})_{i,j=1}^N,$$

where the $a_{ij}$ are defined on a common classical probability space $(\Omega, P)$.

We should formulate this in our algebraic frame of "non-commutative probability spaces". It's instructive to consider the two parts, classical probability spaces and matrices, separately, before we combine them to random matrices.

First, we should write a classical probability space $(\Omega, P)$ in algebraic language. Clearly, we take an algebra of random variables, the only question is which algebra too choose. The best choice for our purposes are the random variables which possess all moments, which we denote by

$$L^{\infty-}(P) := \bigcap_{1 \leq p < \infty} L^p(P).$$

Clearly, the linear functional there is given by taking the expectation

$$E : L^{\infty-}(P) \to \mathbb{C}, \qquad Y \mapsto E[Y] := \int_\Omega Y(\omega) dP(\omega).$$

So a classical probability space $(\Omega, P)$ becomes in our language a non-commutative probability space $(L^{\infty-}(P), E)$ (of course, we are loosing random variables with infinite moments, however, that's good enough for us because usually we are only interested in variables for which all moments exist).

Now let's look on matrices. $N \times N$-matrices $M_N$ come clearly as an algebra, the only question is what state to take. The most canonical one is the trace, and since we need unital functionals we have to normalize it. We will denote the normalized trace on $N \times N$-matrices by tr,

$$\mathrm{tr} : M_N \to \mathbb{C}, \qquad A = (a_{ij})_{i,j=1}^N \mapsto \mathrm{tr}(A) = \frac{1}{N} \sum_{i=1}^N a_{ii}.$$

Maybe we should say a word to justify the choice of the trace. If one is dealing with matrices, then the most important information is contained in their eigenvalues and the most prominent object is the *eigenvalue distribution*. This is a probability measure which puts mass $1/N$ on each eigenvalue (counted with multiplicity) of the matrix. However, note that that's exactly what the trace is doing. Assume we have a symmetric matrix $A$ with eigenvalues $\lambda_1, \ldots, \lambda_N$, then its eigenvalue

distribution is, by definition, the probability measure

$$\nu_A := \frac{1}{N}(\delta_{\lambda_1} + \cdots + \delta_{\lambda_N}).$$

Now note that the unitary invariance of the trace shows that the moments of this measure are exactly the moments of our matrix with respect to the trace,

$$\mathrm{tr}(A^k) = \int_{\mathbb{R}} t^k d\nu_A(t) \qquad \text{for all } k \in \mathbb{N}.$$

This says that the eigenvalue distribution $\nu_A$ is exactly the distribution (which we denoted by $\mu_A$) of the matrix $A$ with respect to the trace. So the trace encodes exactly the right kind of information about our matrices.

Now let us combine the two situations above to random matrices.

**Definition 2.** A non-commutative probability space of $N \times N$-*random matrices* is given by $(M_N \otimes L^{\infty-}(P), \mathrm{tr} \otimes E)$, where $(\Omega, P)$ is a classical probability space. More concretely,

$$A = (a_{ij})_{i,j=1}^N, \qquad \text{with} \qquad a_{ij} \in L^{\infty-}(P)$$

and

$$(\mathrm{tr} \otimes E)(A) = E[\mathrm{tr}(A)] = \frac{1}{N} \sum_{i=1}^N E[a_{ii}].$$

Of course, in this generality there is not much to say about random matrices; for concrete statements we have to specify the classical distribution $P$; i.e., we have to specify the joint distribution of the entries of our matrices.

3.3. **Gaussian random matrices and genus expansion.** A selfadjoint Gaussian random matrix $A = (a_{ij})_{i,j=1}^N$ is a special random matrix where the distribution of the entries is specified as follows:

- the matrix is selfadjoint, $A = A^*$, which means for its entries:
$$a_{ij} = \bar{a}_{ji} \qquad \text{for all } i, j = 1, \ldots, N$$
- apart from this restriction on the entries, we assume that they are independent and complex Gaussians with variance $1/N$.

We can summarize this in the form that the entries $a_{ij}$ $(i, j = 1, \ldots, N)$ form a complex Gaussian family which is determined by the covariance

$$E[a_{ij} a_{kl}] = \frac{1}{N} \delta_{il} \delta_{jk}.$$

Let us try to calculate the moments of such a Gaussian random matrix, with respect to the state

$$\varphi := \mathrm{tr} \otimes E.$$

We have

$$\varphi(A^m) = \frac{1}{N} \sum_{i(1),\ldots,i(m)=1}^{N} E[a_{i(1)i(2)} a_{i(2)i(3)} \cdots a_{i(m)i(1)}].$$

Now we use the fact that the entries of our matrix form a Gaussian family with the covariance as described above, so, by using the Wick formula, we can continue with (where we count modulo $m$, i.e., we put $i(m+1) := i(1)$)

$$\begin{aligned}
\varphi(A^m) &= \frac{1}{N} \sum_{i(1),\ldots,i(m)=1}^{N} \sum_{\pi \in \mathcal{P}_2(m)} \prod_{(r,s)\in\pi} E[a_{i(r)i(r+1)} a_{i(s)i(s+1)}] \\
&= \frac{1}{N} \sum_{i(1),\ldots,i(m)=1}^{N} \sum_{\pi \in \mathcal{P}_2(m)} \prod_{(r,s)\in\pi} \delta_{i(r)i(s+1)} \delta_{i(s)i(r+1)} \frac{1}{N^{m/2}} \\
&= \frac{1}{N^{1+m/2}} \sum_{\pi \in \mathcal{P}_2(m)} \sum_{i(1),\ldots,i(m)=1}^{N} \prod_{(r,s)\in\pi} \delta_{i(r)i(s+1)} \delta_{i(s)i(r+1)}.
\end{aligned}$$

It is now convenient to identify a pairing $\pi \in \mathcal{P}_2(m)$ with a special permutation in $S_m$, just by declaring the blocks of $\pi$ to be cycles; thus $(r,s) \in \pi$ means then $\pi(r) = s$ and $\pi(s) = r$. The advantage of this interpretation becomes apparent from the fact that in this language we can rewrite our last equation as

$$\begin{aligned}
\varphi(A^m) &= \frac{1}{N^{1+m/2}} \sum_{\pi \in \mathcal{P}_2(m)} \sum_{i(1),\ldots,i(m)=1}^{N} \prod_{r=1}^{m} \delta_{i(r)i(\pi(r)+1)} \\
&= \frac{1}{N^{1+m/2}} \sum_{\pi \in \mathcal{P}_2(m)} \sum_{i(1),\ldots,i(m)=1}^{N} \prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))},
\end{aligned}$$

where $\gamma \in S_m$ is the cyclic permutation with one cycle,

$$\gamma = (1, 2, \ldots, m-1, m).$$

If we also identify an $m$-index tuple $i = (i(1), \ldots, i(m))$ with a function $i : \{1, \ldots, m\} \to \{1, \ldots, N\}$, then the meaning of $\prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))}$ is quite obvious, namely it says that the function $i$ must be constant on

the cycles of the permutation $\gamma\pi$ in order to contribute a factor 1, otherwise its contribution will be zero. But in this interpretation

$$\sum_{i(1),\ldots,i(m)=1}^{N} \prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))}$$

is very easy to determine: for each cycle of $\gamma\pi$ we can choose one of the numbers $1,\ldots,N$ for the constant value of $i$ on this orbit, all these choices are independent from each other, which means

$$\sum_{i(1),\ldots,i(m)=1}^{N} \prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))} = N^{\text{number of cycles of } \gamma\pi}.$$

It seems appropriate to have a notation for the number of cycles of a permutation, so let us put

$$\#(\sigma) := \text{number of cycles of } \sigma \qquad (\sigma \in S_m).$$

So we have finally got

$$\varphi(A^m) = \sum_{\pi\in\mathcal{P}_2(m)} N^{\#(\gamma\pi)-1-m/2}.$$

This type of expansion for moments of random matrices is usually called a *genus expansion*, because pairings in $S_m$ can also be identified with surfaces (by gluing the edges of an $m$-gon together according to $\pi$) and then the corresponding exponent of $N$ can be expressed (via Euler's formula) in terms of the genus $g$ of the surface,

$$\#(\gamma\pi) - 1 - m/2 = -2g$$

Let us look at some examples. Clearly, since there are no pairings of an odd number of elements, all odd moments of $A$ are zero. So it is enough to consider the even powers $m = 2k$.

For $m = 2$, the formula just gives a contribution for the pairing $(1, 2) \in S_2$,

$$\varphi(A^2) = 1.$$

There is nothing too exciting about this, essentially we have made our normalization with the factor $1/N$ for the variances of the entries to ensure that $\varphi(A^2)$ is equal to 1 (in particular, does not depend on $N$).

The first non-trivial case is $m = 4$, then we have three pairings, the relevant information about them is contained in the following table.

$$\begin{array}{ccc} \pi & \gamma\pi & \#(\gamma\pi) - 3 \\ (12)(34) & (13)(2)(4) & 0 \\ (13)(24) & (1432) & -2 \\ (14)(23) & (1)(24)(3) & 0 \end{array},$$

so that we have

$$\varphi(A^4) = 2 \cdot N^0 + 1 \cdot N^{-2}$$

For $m = 6, 8, 10$ an inspection of the 15, 105, 945 pairings of six, eight, ten elements yields in the end

$$\begin{aligned} \varphi(A^6) &= 5 \cdot N^0 + 10 \cdot N^{-2} \\ \varphi(A^8) &= 14 \cdot N^0 + 70 \cdot N^{-2} + 21 \cdot N^{-4} \\ \varphi(A^10) &= 42 \cdot N^0 + 420 \cdot N^{-2} + 483 \cdot N^{-4}. \end{aligned}$$

As one might suspect the pairings contributing in leading order $N^0$ are exactly the non-crossing pairings; e.g., in the case $m = 6$ the pairings

$$(12)(34)(56), \quad (12)(36)(45), \quad (14)(23)(56)$$
$$(16)(23)(45), \quad (16)(25)(34).$$

This is true in general. It is quite easy to see that $\#(\gamma\pi)$ can, for a pairing $\pi \in S_{2k}$, at most be $1 + k$, and this upper bound is exactly achieved for non-crossing pairings. In the geometric language of genus expansion, the non-crossing pairings correspond to genus zero or planar situations.

But this tells us that, although the moments of a $N \times N$ Gaussian random matrix for fixed $N$ are quite involved, in the limit $N \to \infty$ they become much simpler and converge to something which we understand quite well, namely the Catalan numbers $c_k$,

$$\lim_{N\to\infty} \varphi(A^{2k}) = c_k = \frac{1}{k+1}\binom{2k}{k}.$$

We should remark here that the subleading contributions in the genus expansion are much harder to understand than the leading order. There are no nice explicit formulas for the coefficients of $N^{-2g}$. (However, there are some recurrence relations between them, going back to a well-known paper of Harer and Zagier, 1986.)

This limit moments of our Gaussian random matrices is something which we know quite well from free probability theory – these are the moments of the most important element in free probability, the semi-circular element. So we see that Gaussian random matrices realize, at least asymptotically in the limit $N \to \infty$, semi-circular elements (e.g., the sum of creation and annihilation operator on full Fock space).

3.4. **Asymptotic freeness for several independent Gaussian random matrices.** The fact that the eigenvalue distribution of Gaussian random matrices converges in the limit $N \to \infty$ to the semi-circle distribution is one of the basic facts in random matrix theory, it was proved by Wigner in 1955, and is accordingly usually termed *Wigner's semi-circle law*. Thus the semi-circle distribution appeared as an interesting object long before semi-circular elements in free probability were considered. This raised the question whether it is just a coincidence that Gaussian random matrices in the limit $N \to \infty$ and the sum of creation and annihilation operators on full Fock spaces have the same distribution or whether there is some deeper connection. Of course, our main interest is in the question whether there is also some freeness around (at least asymptotically) for random matrices. That this is indeed the case was one of the most fruitful discoveries of Voiculescu in free probability theory.

In order to have freeness we should of course consider two random matrices. Let us try the simplest case, by taking two Gaussian random matrices

$$A^{(1)} = (a_{ij}^{(1)})_{i,j=1}^{N}, \qquad A^{(2)} = (a_{ij}^{(2)})_{i,j=1}^{N}.$$

Of course, we must also specify the relation between them, i.e., we must prescribe the joint distribution of the whole family

$$a_{11}^{(1)}, \ldots, a_{NN}^{(1)}, a_{11}^{(2)}, \ldots, a_{NN}^{(2)}.$$

Again we stick to the simplest possible case and assume that all entries of $A^{(1)}$ are independent from all entries of $A^{(2)}$; i.e., we consider now two independent Gaussian random matrices. This is the most canonical thing one can do from a random matrix perspective. To put it down more formally (and in order to use our Wick formula), the collection of all entries of our two matrices forms a Gaussian family with covariance

$$E[a_{ij}^{(r)} a_{kl}^{(p)}] = \frac{1}{N} \delta_{il} \delta_{jk} \delta_{rp}$$

Let us now see whether we can extend the above calculation to this situation. And actually, it turns out that we can just repeat the above

calculation with putting superindices $p(1),\ldots,p(m) \in \{1,2\}$ to our matrices. The main arguments are not affected by this.

$$\varphi(A^{(p(1))}\cdots A^{(p(m))}) = \frac{1}{N}\sum_{i(1),\ldots,i(m)=1}^{N} E[a^{(p(1))}_{i(1)i(2)}a^{(p(2))}_{i(2)i(3)}\cdots a^{(p(m))}_{i(m)i(1)}]$$

$$= \frac{1}{N}\sum_{i(1),\ldots,i(m)=1}^{N}\sum_{\pi\in\mathcal{P}_2(m)}\prod_{(r,s)\in\pi} E[a^{(p(r))}_{i(r)i(r+1)}a^{(p(s))}_{i(s)i(s+1)}]$$

$$= \frac{1}{N}\sum_{i(1),\ldots,i(m)=1}^{N}\sum_{\pi\in\mathcal{P}_2(m)}\prod_{(r,s)\in\pi} \delta_{i(r)i(s+1)}\delta_{i(s)i(r+1)}\delta_{p(r)p(s)}\frac{1}{N^{m/2}}$$

$$= \frac{1}{N^{1+m/2}}\sum_{\pi\in\mathcal{P}_2(m)}\sum_{i(1),\ldots,i(m)=1}^{N}\prod_{r=1}^{m}\delta_{i(r)i(\gamma\pi(r))}\delta_{p(r)p(\pi(r))}.$$

The only difference to the situation of one matrix is now the extra factor $\prod_{r=1}^{m}\delta_{p(r)p(\pi(r))}$, which just says that we have an extra condition on our pairings $\pi$, namely they must pair the same matrices, i.e., no block of $\pi$ is allowed to pair $A^{(1)}$ with $A^{(2)}$. If $\pi$ has this property then its contribution will be as before, otherwise it will be zero. Thus we might, for $p=(p(1),\ldots,p(m))$, introduce the notation

$$\mathcal{P}_2^{(p)}(m) := \{\pi\in\mathcal{P}_2(m) \mid p(\pi(r))=p(r) \text{ for all } r=1,\ldots,m\}.$$

Then we can write the final conclusion of our calculation as

$$\varphi(A^{(p(1))}\cdots A^{(p(m))}) = \sum_{\pi\in\mathcal{P}_2^{(p)}(m)} N^{\#(\gamma\pi)-1-m/2}.$$

Here are a few examples for $m=4$.

$$\varphi(A^{(1)}A^{(1)}A^{(1)}A^{(1)}) = 2\cdot N^0 + 1\cdot N^{-2}$$
$$\varphi(A^{(1)}A^{(1)}A^{(2)}A^{(2)}) = 1\cdot N^0 + 0\cdot N^{-2}$$
$$\varphi(A^{(1)}A^{(2)}A^{(1)}A^{(2)}) = 0\cdot N^0 + 1\cdot N^{-2}$$

As before, the leading term for $N\to\infty$, is given by contributions from non-crossing pairings, but now these non-crossing pairings must connect an $A^{(1)}$ with an $A^{(1)}$ and an $A^{(2)}$ with an $A^{(2)}$. But this is exactly the rule for calculating mixed moments in two semi-circular elements which are free. Thus we see that we indeed have asymptotic freeness between two independent Gaussian random matrices. Of course, the same is true if we consider $n$ independent Gaussian random matrices, they are becoming asymptotically free. Another way of saying this is that we cannot only realize one sum $\omega$ of creation

and annihilation operator by a Gaussian random matrix, but we can also realize a genuine non-commutative situation of $n$ such operators $\omega(f_1), \ldots, \omega(f_n)$ by $n$ independent Gaussian random variables. Since $n$ free semi-circulars generate the free group factor $L(\mathbb{F}_n)$, we have now suddenly an asymptotic realization of $L(\mathbb{F}_n)$ by $n$ independent Gaussian random matrices. This is a quite non-standard realization which provides a quite different point of view on the free group factors. It remains to see whether this realization allows to gain new insight into the free group factors. Lecture 4 will mostly be about the harvest of this approach.

But for the moment we want to continue our investigation on the general relation between random matrices and freeness and, encouraged by the above result, we are getting more ambitious in this direction.

3.5. **Asymptotic freeness between Gaussian random matrices and constant matrices.** The above generalization of Voiculescu of Wigner's semi-circle law is on one side really a huge step; we do not only find the semi-circular distribution in random matrices, but the concept of freeness itself shows up very canonically for random matrices. However, on the negative side, one might argue that in the situation considered above we find freeness only for a very restricted class of distributions, namely for semi-circulars. In classical probability theory, this would be comparable to saying that we understand the concept of independence for Gaussian families. Of course, this is only a very restricted version and we should aim at finding more general appearances of freeness in the random matrix world.

Here is the next step in this direction. Instead of looking on the relation between two Gaussian random matrices we replace now one of them by a "constant" matrix. This just means ordinary matrices $M_N$ without any randomness, and the state there is just given by the trace tr. Of course, we only expect asymptotic freeness in the limit $N \to \infty$, so what we really are looking at is a sequence of matrices $D_N$ whose distribution, with respect to tr, converges asymptotically. Thus we assume the existence of all limits

$$\lim_{N\to\infty} \mathrm{tr}(D_N^m) \qquad (m \in \mathbb{N}).$$

Note that we have a large freedom of prescribing the wanted moments in the limit. E.g., we can take diagonal matrices for the $D_N$ and then we can approximate any fixed, let's say compactly supported, probability measure on $\mathbb{R}$ by suitably choosen matrices.

As for the Gaussian random matrices we will usually supress the index $N$ at our constant matrices to lighten the notation. But keep in

mind that we are talking about sequences of $N \times N$-matrices and we want to take the limit $N \to \infty$ in the end.

Let us now see how far we can go with our above calculations in such a situation. What we could like to understand are mixed moments in a Gaussian random matrix $A$ and a constant matrix $D$. We can bring this always in the form

$$\varphi(D^{(q(1))} A D^{(q(2))} A \cdots D^{(q(m))} A)$$

where $D^{q(i)}$ are some powers of the matrix $D$. It is notationally convenient to introduce new names for these powers and thus go formally over from one matrix $D$ to a set of several constant matrices (and assuming the existence of all their joint moments), because then it suffices to look on moments which are alternating in $A$ and $D$'s. A power of a constant matrix is just another constant matrix, so we are not loosing anything. Of course, for our Gaussian random matrix $A$ we prefer to keep $A$ because the power of a Gaussian random matrix is not a Gaussian random matrix anymore, but some more complicated random matrix.

Let us now do the calculation of such alternating moments in $D$'s and $A$'s.

$$\begin{aligned}
&\varphi(D^{(q(1))} A \cdots D^{(q(m))} A) \\
&= \frac{1}{N} \sum_{\substack{i(1),\ldots,i(m),\\ j(1),\ldots,j(m)=1}}^{N} E[d^{(q(1))}_{j(1)i(1)} a_{i(1)j(2)} d^{(q(2))}_{j(2)i(2)} a_{i(2)j(3)} \cdots d^{(q(m))}_{j(m)i(m)} a_{i(m)j(1)}] \\
&= \frac{1}{N} \sum_{\substack{i(1),\ldots,i(m),\\ j(1),\ldots,j(m)=1}}^{N} E[a_{i(1)j(2)} a_{i(2)j(3)} \cdots a_{i(m)j(1)}] \cdot d^{(q(1))}_{j(1)i(1)} d^{(q(2))}_{j(2)i(2)} \cdots d^{(q(m))}_{j(m)i(m)} \\
&= \frac{1}{N} \sum_{\substack{i(1),\ldots,i(m),\\ j(1),\ldots,j(m)=1}}^{N} \sum_{\pi \in \mathcal{P}_2(m)} \prod_{(r,s) \in \pi} E[a_{i(r)j(r+1)} a_{i(s)j(s+1)}] \cdot d^{(q(1))}_{j(1)i(1)} d^{(q(2))}_{j(2)i(2)} \cdots d^{(q(m))}_{j(m)i(m)} \\
&= \frac{1}{N} \sum_{\substack{i(1),\ldots,i(m),\\ j(1),\ldots,j(m)=1}}^{N} \sum_{\pi \in \mathcal{P}_2(m)} \prod_{(r,s) \in \pi} \delta_{i(r)j(s+1)} \delta_{i(s)j(r+1)} \frac{1}{N^{m/2}} \cdot d^{(q(1))}_{j(1)i(1)} d^{(q(2))}_{j(2)i(2)} \cdots d^{(q(m))}_{j(m)i(m)}
\end{aligned}$$

Again, we identify a $\pi \in \mathcal{P}_2(m)$ with a permutation in $S_m$ and then it remains to understand, for such a fixed $\pi$, the expression

$$\sum_{\substack{i(1),\ldots,i(m),\\ j(1),\ldots,j(m)=1}}^{N} \prod_{(r,s)\in\pi} \delta_{i(r)j(s+1)}\delta_{i(s)j(r+1)} \cdot d^{(q(1))}_{j(1)i(1)} d^{(q(2))}_{j(2)i(2)} \cdots d^{(q(m))}_{j(m)i(m)} =$$

$$= \sum_{\substack{i(1),\ldots,i(m),\\ j(1),\ldots,j(m)=1}}^{N} \prod_{r=1}^{m} \delta_{i(r)j(\gamma\pi(r))} \cdot d^{(q(1))}_{j(1)i(1)} d^{(q(2))}_{j(2)i(2)} \cdots d^{(q(m))}_{j(m)i(m)}$$

$$= \sum_{j(1),\ldots,j(m)=1}^{N} d^{(q(1))}_{j(1)j(\gamma\pi(1))} d^{(q(2))}_{j(2)j(\gamma\pi(2))} \cdots d^{(q(m))}_{j(m)j(\gamma\pi(m))}$$

One realizes that this factors into a product of traces; one trace for each cycle of $\gamma\pi$.

To see a concrete example, consider $\sigma = (1,3,6)(4)(2,5) \in S_6$. Then one has

$$\sum_{j(1),\ldots,j(6)=1}^{N} d^{(1)}_{j(1)j(\sigma(1))} d^{(2)}_{j(2)j(\sigma(2))} d^{(3)}_{j(3)j(\sigma(3))} d^{(4)}_{j(4)j(\sigma(4))} d^{(5)}_{j(5)j(\sigma(5))} d^{(6))}_{j(6)j(\sigma(6}$$

$$= \sum_{j(1),\ldots,j(6)=1}^{N} d^{(1)}_{j(1)j(3)} d^{(2)}_{j(2)j(5)} d^{(3)}_{j(3)j(6)} d^{(4)}_{j(4)j(4)} d^{(5)}_{j(5)j(2)} d^{(6)}_{j(6)j(1)}$$

$$= \sum_{j(1),\ldots,j(6)=1}^{N} d^{(1)}_{j(1)j(3)} d^{(3)}_{j(3)j(6)} d^{(6)}_{j(6)j(1)} \cdot d^{(2)}_{j(2)j(5)} d^{(5)}_{j(5)j(2)} \cdot d^{(4)}_{j(4)j(4)}$$

$$= \mathrm{Tr}[D^{(1)}D^{(3)}D^{(6)}] \cdot \mathrm{Tr}[D^{(2)}D^{(5)}] \cdot \mathrm{Tr}[D^{(4)}]$$

$$= N^3 \mathrm{tr}[D^{(1)}D^{(3)}D^{(6)}] \cdot \mathrm{tr}[D^{(2)}D^{(5)}] \cdot \mathrm{tr}[D^{(4)}],$$

which is exactly a product of traces along the cycles of $\sigma$. (Note that we denoted by Tr the unnormalized trace.) It seems, we need again some compact notation for this kind of product. Let us put, for $\sigma \in S_m$,

$$\mathrm{tr}_\sigma[D^{(1)},\ldots,D^{(m)}] := \sum_{j(1),\ldots,j(m)=1}^{N} d^{(1)}_{j(1)j(\sigma(1))} \cdots d^{(m)}_{j(m)j(\sigma(m))}$$

$$= \prod_{\text{cycles of } \sigma} \mathrm{tr}\Big( \prod_{\text{along the cycle}} D^{(\ldots)} \Big).$$

Let us summarize our calculation.

$$\varphi(D^{(q(1))}A \cdots D^{(q(m))}A) = \sum_{\pi\in\mathcal{P}_2(m)} \mathrm{tr}_{\gamma\pi}[D^{(q(1))},\ldots,D^{(q(m))}] \cdot N^{\#(\gamma\pi)-1-m/2}.$$

Of course, we can do the same for several instead of one Gaussian random matrix, the only effect of this is to restrict the sum over $\pi$ to pairings which respect the "colour" of the matrices.

$$\varphi(D^{(q(1))}A^{(p(1))}\cdots D^{(q(m))}A^{(p(m))}) = \sum_{\pi\in\mathcal{P}_2^{(p)}(m)} \mathrm{tr}_{\gamma\pi}[D^{(q(1))},\ldots,D^{(q(m))}]\cdot N^{\#(\gamma\pi)-1-m/2}.$$

Now let's look on the asymptotic structure of this formula. By our assumption $\mathrm{tr}_{\gamma\pi}[D^{(q(1))},\ldots,D^{(q(m))}]$ has a limit, and the factor $N^{\#(\gamma\pi)-1-m/2}$ surpreses all crossing pairings. If we denote the limit of our moments by $\psi$, then we have

$$\psi(D^{(q(1))}A^{(p(1))}\cdots D^{(q(m))}A^{(p(m))}) = \sum_{\pi\in NC_2^{(p)}(m)} \psi_{\gamma\pi}[D^{(q(1))},\ldots,D^{(q(m))}].$$

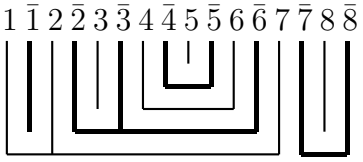
This resembles our formula from the last lecture for alternating moments in two free families of random variables, if one of them is a semi-circular family. The only difference is that there we have the inverse complement $K^{-1}(\pi)$ of $\pi$ whereas here we have $\gamma\pi$. However, this is exactly the same. For a non-crossing partition $\pi$, identified with an element in $S_n$, the block structure of its complement is given by the cycles of $\gamma^{-1}\pi$. Let us just check this on an example. Consider the partition

$$\pi := (1,2,7),(3),(4,6),(5),(8) \in NC(8).$$

One has

$$\gamma^{-1}\pi = (1)(2,6,3)(4,5)(7,8),$$

which gives us, by changing the cycles to blocks, the complement $K(\pi)$, as can be seen from the graphical representation:




Let us collect what we have observed about asymptotic freeness of random matrices in the next theorem.

**Theorem 1.** *Let $A_N$ be a Gaussian $N\times N$-random matrix and $D_N$ a constant (non-probabilistic) matrix, such that all limiting moments*

$$\lim_{N\to\infty} tr(D^k) \qquad (k\in\mathbb{N})$$

*exist. Then $A_N$ and $D_N$ are asymptotically free.*

Of course, the meaning of asymptotic freeness should be clear, namely all mixed moments between $A_N$ and $D_N$ compute in the limit $N \to \infty$ as for free random variable. In order to formalize this a bit more, one can introduce a notion of *convergence (in distribution)*, saying that random variables $a_1^{(N)}, \ldots, a_m^{(N)}$ converge to random variables $a_1, \ldots, a_m$ if each moment in the $a^{(N)}$'s converges to the corresponding moment in the $a$'s. We denote this formally by

$$a_1^{(N)}, \ldots, a_m^{(N)} \to a_1, \ldots, a_m.$$

(This is the analogue of the classical notion of "convergence in law"; as there the variables for each $N$ and in the limit might live on different probability spaces.)

With this notion, we can rephrase the above theorem in the following form: Assume that the matrices $D_N$ converge to a random variable $d$, then the matrices $A_N, D_N$ converge to a pair $s, d$, where $s$ is a semi-circular element and $s$ and $d$ are free. Let us phrase the case of more matrices directly in this language.

**Theorem 2.** *Let $A_N^{(1)}, \ldots, A_N^{(p)}$ be $p$ independent Gaussian random matrices and let $D_N^{(1)}, \ldots, D_N^{(q)}$ be $q$ constant matrices which converge for $N \to \infty$ to random variables $d_1, \ldots, d_q$,i.e.,*

$$D_N^{(1)}, \ldots, D_N^{(q)} \to d_1, \ldots, d_q.$$

*Then*

$$A_N^{(1)}, \ldots, A_N^{(p)}, D_N^{(1)}, \ldots, D_N^{(q)} \to s_1, \ldots, s_p, d_1, \ldots, d_q,$$

*where each $s_i$ is a semi-circular element and where $s_1, \ldots, s_p, \{d_1, \ldots, d_q\}$ are free.*

These considerations show that random matrices allow to realize freeness between semi-circular and anything. Of course, the final question is whether we can reach the ultimate level and realize freeness between anything and anything else in random matrices. That this is indeed the case relies on having similar asymptotic freeness statements for unitary random matrices instead of Gaussian ones.

3.6. **Asymptotic freeness between Haar unitary random matrices and constant matrices.** Another important random matrix ensemble is given by Haar unitary random matrices. We equip the compact group $\mathcal{U}(N)$ of unitary $N \times N$-matrices with its Haar probability measure and accordingly distributed random matrices we shall call *Haar distributed unitary random matrices*. Thus the expectation

$E$ over this ensemble is given by integrating with respect to the Haar measure.

If we are going to imitate for this case the calculation we did for Gaussian random matrices, then we need the expectations of products in entries of such a unitary random matrix $U$. In contrast to the Gaussian case, this is now not explicitly given. In particular, entries of $U$ won't in general be independent. However, one expects that the unitary condition $U^*U = 1 = UU^*$ and the invariance of the Haar measure under multiplication from right or from left with an arbitrary unitary matrix should allow to determine the mixed moments of the entries of $U$. This is indeed the case; the calculation, however, is not trivial and we only describe in the following the final result, which has the flavour of the Wick formula for the Gaussian case.

The expectation of products of entries of Haar distributed unitary random matrices can be described in terms of a special function on the permutation group. Since such considerations go back to Weingarten, we call this function the Weingarten function and denotes it by Wg. (However, it was Feng Xu who revived these ideas and made the connection with free probability.) In the following we just recall the relevant information about this Weingarten function.

We use the following definition of the Weingarten function. For $\pi \in S_n$ and $N \geq n$ we put

$$\mathrm{Wg}(N, \pi) = E\big(u_{11} \cdots u_{nn} \overline{u}_{1\pi(1)} \cdots \overline{u}_{n\pi(n)}\big),$$

where $U = (u_{ij})_{i,j=1}^N$ is a Haar distributed unitary $N \times N$-random matrix. This $\mathrm{Wg}(N, \pi)$ depends on $\pi$ only through its conjugacy class. General matrix integrals over the unitary groups can then be calculated as follows:

$$E\big(u_{i'_1 j'_1} \cdots u_{i'_n j'_n} \overline{u}_{i_1 j_1} \cdots \overline{u}_{i_n j_n}\big) = \sum_{\alpha,\beta \in S_n} \delta_{i_1 i'_{\alpha(1)}} \cdots \delta_{i_n i'_{\alpha(n)}} \delta_{j_1 j'_{\beta(1)}} \cdots \delta_{j_n j'_{\beta(n)}} \mathrm{Wg}(N, \beta\alpha^{-1}).$$

(Note that corresponding integrals for which the number of $u$'s and $\bar{u}$'s is different vanish, by the invariance of such an expression under the replacement $U \mapsto \lambda U$, where $\lambda \in \mathbb{C}$ with $|\lambda| = 1$.) The Weingarten function is a quite complicated object. For our purposes, however, only the behaviour of leading orders in $N$ of $\mathrm{Wg}(N, \pi)$ is important. One knows that the leading order in $1/N$ is given by $2n - \#(\pi)$ ($\pi \in S_n$) and increases in steps of 2.

Using the above formula for expectations of products of entries of $U$ and the information on the leading order of the Weingarten function

one can show, in a similar way as for the Gaussian random matrices, that Haar unitary random matrices and constant matrices are asymptotically free. In this context one should surely first think about what the limit distribution of a Haar unitary random matrix is. This is – not only in the limit, but even for each $N$ – a Haar unitary element in the following sense.

**Definition 3.** Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space with $\mathcal{A}$ being a $*$-algebra. An element $u \in \mathcal{A}$ is called *Haar unitary* if it is unitary and if

$$\varphi(u^k) = 0 \qquad \text{for all } k \in \mathbb{Z}\backslash\{0\}.$$

Clearly, the distribution $\mu_u$ of a Haar unitary is the uniform distribution on the circle of radius 1. Note furthermore that the unitary operators generating the free group factors in the defining (left regular) representation of $L(\mathbb{F}_n)$ are Haar unitaries. Thus we can say that our original realization of $L(\mathbb{F}_n)$ is in terms of $n$ free Haar unitaries.

Let me now state the asymptotic freeness result for Haar unitary random matrices.

**Theorem 3.** *Let $U_N^{(1)}, \ldots, U_N^{(p)}$ be $p$ independent Haar unitary random matrices and let $D_N^{(1)}, \ldots, D_N^{(q)}$ be $q$ constant matrices which converge for $N \to \infty$ to random variables $d_1, \ldots, d_q$, i.e.,*

$$D_N^{(1)}, \ldots, D_N^{(q)} \to d_1, \ldots, d_q.$$

*Then*

$$U_N^{(1)}, U_N^{(1)*}, \ldots, U_N^{(p)}, U_N^{(p)*}, D_N^{(1)}, \ldots, D_N^{(q)} \to u_1, u_1^*, \ldots, u_p, u_p^*, d_1, \ldots, d_q,$$

*where each $u_i$ is a Haar unitary element and where $u_1, \ldots, u_p, \{d_1, \ldots, d_q\}$ are $*$-free.*

3.7. **Making matrices free by random rotations.** This does not necessarily look like an improvement over the Gaussian case, but this allows us to realize freeness between quite arbitrary distributions by the following simple observation (to prove it, just use the definition of freeness).

**Proposition 1.** *Assume $u$ is a Haar unitary and $u$ is $*$-free from $\{a, b\}$. Then $a$ and $ubu^*$ are also free.*

This yields finally the following general recipe for realizing free distributions by random matrices.

**Theorem 4.** *Let $A_N$ and $B_N$ be two sequences of constant $N \times N$-matrices with limiting distributions*

$$A_N \to a, \qquad B_N \to b.$$

*Let $U_N$ be a Haar unitary $N \times N$-random matrix. Then we have*

$$A_N, U_N B_N U_N^* \to a, b,$$

*where $a$ and $b$ are free.*

Note that conjugating by $U_N$ does not change the distribution of $B_N$. The effect of this "random rotation" is that any possible relation between $A_N$ and $B_N$ is destroyed, and these two matrices are brought in "generic" position. This gives maybe the most concrete picture for freeness. $N \times N$-matrices are asymptotically free, if their eigenspaces are in generic position, have no definite relation between them.

The idea of bringing matrices in a generic, free position by random unitary conjugation also works more general for *random* matrices if we have in addition independence between the involved ensembles.

**Theorem 5.** *Let $U$ be a sequence of Haar distributed unitary $N \times N$-random matrices. Suppose that $A_1, \ldots, A_s$ and $B_1, \ldots, B_t$ are sequences of $N \times N$-random matrices each of which has a limit distribution. Furthermore, assume that $\{A_1, \ldots, A_s\}$, $\{B_1, \ldots, B_t\}$, and $\{U\}$ are independent. Then the sequences $\{A_1, \ldots, A_s\}$ and $\{UB_1U^*, \ldots, UB_tU^*\}$ are asymptotically free.*

3.8. **Using free probability random matrix calculations.** Instead of using random matrices for modelling operator algebras (as we will soon), one can also revert the direction and use our machinery for calculating moments of free variables in order to calculate eigenvalue distributions of random matrices in the limit $N \to \infty$. Such calculations are one important part of random matrix theory (in particular, in applications to engineering) and free probability theory has added quite a bit to this. Here are some examples what free probability theory can do.

- If we have two random matrices which are free (e.g., of the form $A, UBU^*$ for a Haar unitary random matrix $U$), then we can use our $R$-transform or $S$-transform machinery for calculating the asymptotic eigenvalue distribution of the sum or of the product of these matrices.
- There exist also formulas (due to Nica and myself) describing the distribution of the commutator $ab - ba$ of two free random variables. This can be used for calculating the distribution of the commutator of two random matrices in generic position.
- We have also seen in the last lecture that we have a nice description of the free compression of random variables. However, this has also a very concrete meaning for random matrices. If we

have a randomly rotated matrix $UAU^*$, then this is asymptotically free from diagonal projections, and taking the compression with these diagonal projections just means to take corners of the original matrix. Thus free probability tells us in this case how to calculate the distribution of corners of matrices out of the distribution of the whole matrix.

Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
*E-mail address*: speicher@mast.queensu.ca